Preface

In 2001 the United Kingdom suffered a major outbreak of foot-and-mouth disease in livestock. The scientific experts assembled to advise the government included several groups of mathematical modelers. Within weeks of the first case report, modelers had developed systems of equations describing the farm-to-farm spread of infection, and were using them to help devise measures for controlling the outbreak (Ferguson et al. 2001a, b; Keeling et al. 2001). Adapting previous models, Ferguson et al. (2001) classified farms according to disease state: disease-free, infected, and vaccinated or slaughtered, with the infected class subdivided into several stages to represent the degree of infection and whether or not the disease had been detected. Equations were developed to describe how farms become infected, experience a local outbreak, and pass the disease to other farms, with model parameters estimated from the recorded dates and locations of all reported infections. Because the model is based on equations representing the process of disease transmission and spread, it could be modified to incorporate potential control measures and then used to predict their consequences. It was concluded that immediate slaughter of infected herds would slow disease spread but not halt it, and more aggressive measures were needed. Based in part on those analyses, the government instituted a policy of rapidly culling all farms in a ring surrounding a farm where infection is detected. These policies were successful: the disease soon started to decline and the outbreak collapsed within a few months. Detailed modeling and data analyses later confirmed that the control measures had greatly reduced the size of the outbreak, but also suggested that earlier application of the strict control measures that were eventually imposed could have reduced the outbreak by about half (Keeling et al. 2001).

The 2001 foot-and-mouth outbreak is a particularly dramatic example of how models, written in the language of mathematics and solved on the computer, are

now used throughout biology. Experimental biology is reductionist, disclosing how life works one piece at a time. But at all levels life is *complex*, meaning that the properties of cells, organisms, populations, and ecosystems reflect the interactions among multiple components. In virtually all areas of biology, models provide the means for putting the pieces together, for example,

- to understand the consequences of multiple, interacting signal transduction pathways within cells and complex networks of interaction among genes;
- to explain the development of irregular heart behaviors such as fibrillation, and to improve methods for treating them;
- to predict which species are at greatest risk of extinction, to identify effective measures for their preservation, and to predict how ecosystems will respond to the changes in climate and massive loss of species now in progress.

In the past, biologists have generally been uninterested in mathematics, but that attitude is changing rapidly. With advances in many technologies, our intuition is no longer sufficient to analyze and interpret data. Quantitative models and methods are needed, sometimes because data are so extensive that automated methods are needed to examine them in detail (e.g., genome sequences) and sometimes because the pheomena are sufficiently complex that computer simulation is the only way to derive predictions from a model. We believe that biology students need to connect what they learn in a required calculus course to biology, and that mathematics students are greatly enriched when they see mathematics connected to real-world problems in a more tangible way than is taught in mathematics courses.

These connections come through models, but there are some complex issues that must be addressed to develop an appreciation for them. What are the principles used in developing a model? What simplifications are made to make a model tractable? What measurements are used to estimate model parameters and validate model predictions? How does mathematics help us analyze models, either by hand or through computer algorithms? How robust are the predictions of a model to uncertainties in the structure of the model, errors in our measurements, stochastic fluctuations of the environment, and the individual variability of living organisms? These are all difficult questions, and they deserve sustained attention as students learn the art of modeling biological systems.

This text is our attempt to address the lack of curricular materials designed to introduce undergraduate students to dynamic models of biological systems. It is based on a course that we have taught to a diverse audience of students majoring in the biological sciences, computer science, mathematics, engineering, and the physical sciences. We have adopted the following strategies in writing the text:

- We make no attempt to be comprehensive. Most chapters follow the "business school" model of focusing on a selective set of case studies that are treated in depth.

- Examples have been chosen to span the breadth of major areas within biology and to illustrate diverse types of mathematical models. Models of processes within cells, organisms, and populations are all based on the same principles. At this level, an integrative approach that does not divide biology along these boundaries seems appropriate. The models that we study are either ones that have stood the test of time or recent ones that are especially prominent now. Similarly, some of our case studies are classics, and others are very recent, but all of them are compelling to us and exemplify the reasons why we chose to work in this field.

- We restrict attention to dynamic models. This is a reflection of our personal interests and expertise. Ideally this course would be accompanied by a second semester devoted to genomics, bioinformatics, and structural biology where the mathematical theory is drawn primarily from combinatorics, probability, statistics, and optimization.

- We rely on high-level computer languages such as Matlab or R to implement models and perform calculations. This minimizes the user effort required to implement models, run simulations, and perform mathematical analysis, and allows us to quickly proceed to discussion of modeling results. The primary question we ask is what we learn about the biology from the models.

- We introduce the minimal amount of mathematics needed to understand the modeling results. We hope that students taking the course will be inspired to learn about the mathematics we introduce by studying the material in more depth in mathematics courses — but we make no pretense of discussing more than basic theoretical concepts. In contrast, biomathematics courses often select mathematics that is needed to study models in certain domains and teach this material to students who will engage in modeling within this domain. We think that approach is problematic in that students may need a deeper understanding of the mathematics than is typically covered in such courses.

- We discuss the modeling process itself. The first and last chapters of the book discuss the nature of models and give our advice about how to construct and use models. Intermediate chapters intersperse comments intended to place the models into a broad context, with citations to key references where the student can learn more.

- We expect students to read original papers from the biology literature. The published literature plays very different roles in biology and mathematics, and this influences our use of literature in the course. The mathematics literature records and systematizes theories that an individual is expected to personally reconstruct for full understanding. Biology literature describes experiments or observations and their outcomes, along with the hypotheses that the experiments were designed to test and the interpretation of the results. Critical evaluation of how experimental results are interpreted is an important part of the subject. Since the purpose of models is to address biological questions, we think that it is important for students—even at the level of this course—to engage in such critical thinking.

We require students to do a course project. Their assignment is to choose a published research paper using a dynamic model, implement the model, reproduce the modeling results in the paper, and then take one more step. By this we mean that they should formulate a question that was not fully addressed in the paper, modify or extend the model to do so, and then present their findings. This is a difficult assignment but we have been gratified by the papers that the students have written. Much more than the homework exercises, it prompts students to think creatively and critically about science in a setting in which answers are not absolutely right or wrong.

We think computer exercises are essential so that students will have personal experience as modelers and model-users. We use the first three or four computer lab sessions to teach the rudiments of programming and then add a bit each week as new types of models are encountered—how to compute eigenvalues, generate random numbers, solve differential equations, and so on. Our experience is that undergraduate biology students can and do learn how to do real scientific programming. Instructional materials and suggested exercises for computer labs can be downloaded at no cost from this book's web page at Princeton University Press (http://www.pupress.princeton.edu/titles/8124.html).

There is more in this book than can be covered in a single semester, allowing some flexibility as to the choice of topics and the mathematical level of the course. Here are a few suggestions:

1. Because we do not divide biology along disciplinary lines, premed students get a substantial dose of population biology before we get to neurotransmitters—and they have to pay attention, because we use population models to introduce some essential mathematics for everything that follows. Therefore we strongly recommend covering at least a good fraction of Chapter 1, so that students have a taste of the variety of applications to come.

2. Chapters 2 and 3 and Chapters 4 and 6 are paired in the sense that models of the same type are used in very different applications. The second of each pair can therefore be skipped without losing anything essential for later chapters. With a bit more effort only the second can be covered, with material from the first brought in as needed.

3. Chapters 5, 7, 8, and 9 are independent of each other and are not required for reading any other chapters—feel free to pick and choose, or to cover some of each (as we generally do).

4. Chapters 3 and 7 are the only ones where probability theory plays any significant role. A course on "deterministic models" can be constructed by omitting those chapters, and the sections in chapters 2 and 9 on stochastic models.

Acknowledgments

We would like to thank first our families for their support and patience while we wrote this book.

We are grateful to Cornell University for giving us the opportunity to teach our interdisciplinary class, for providing adequate teaching assistant support, and for a sabbatical leave to Steve Ellner while we completed the book. Thanks also to the Cornell Laboratory of Ornithology for hosting the sabbatical.

The decision to turn our class into a book was inspired by our colleague Rick Durrett's ability to write a book every few years while remaining prolific as a researcher, and by his explanation of how it's done ("do your lecture notes in TeX")—if only it were really that easy (but maybe it is for Rick). We also thank Sam Elworthy at Princeton University Press for encouraging us to start and finish the writing process and Jennifer Slater for her careful copyediting. Jonathan Rowell contributed notes on linear algebra that were the basis for parts of Chapter 2, and Section 6.6 on within-host dynamics of HIV is based on a guest lecture and lecture notes by Laura Jones. The National Science Foundation supported Carla Martin as a VIGRE Fellow to help collect materials for the course the first time we taught it, and a number of the exercises were written or improved by her co-TA Daniel Fink.

The students in the 2002 and 2004 renditions of our course were great. Their enthusiasm was infectious and their close reading of the notes helped us eliminate some mistakes. (Please tell us about any more that you discover!) We have also benefited greatly from comments on draft chapters by Hal Caswell, Leon Glass, Stefan Hames, M. Henry H. Stevens, and several anonymous reviewers. Special thanks to Carole Hom and to Mason Porter and Todd Stokes who field-tested draft chapters in courses at UC Davis and Georgia Tech, respectively, and gave us extensive feedback.

Textbooks do not emerge from the void. We owe a debt to previous authors and colleagues from whom we have learned (and probably "borrowed") a great deal, including Leah Edelstein-Keshet, Lee Segel, Simon Levin, Doug Nychka, Ron Gallant, Simon Wood, Harvey Gold, Charles Hall, Bertil Hille, James Murray, and Art Winfree.

Art and Lee, we miss you. Perhaps some day modeling will help to cure cancer.